



Training for Variable and Innovative Behavior

Karen Pryor and Sheila Chase
Hunter College of the City University of New York, USA

This paper provides a summary of a 1969 report (Pryor, Haag, & O'Reilly) of the spontaneous emergence of innovative behavior of a dolphin¹, a replication of this event through training in another dolphin, and the effect this work has had on current animal training technology. This paper provides a review of laboratory based research in support of some of the procedures found effective in modern animal training in developing innovative behavior, specifically use of the conditioned reinforcer to mark a behavior, differential reinforcement of variability, and intentional use of positive reinforcement procedures. The authors describe specific processes for establishing innovative skills, practical applications presently in use with animals, consequent human and animal welfare benefits, and suggestions for further research.

An Experiment in Training Novel Behavior

The unexpected discovery in the early 1960s of spontaneous innovation by a mature female rough-toothed dolphin (*Steno bredanensis*) named Malia was an outcome of a demonstration to audiences at Sea Life Park, an oceanarium in Hawaii. Trainers intended to show audiences the first step in establishing a new, trained behavior, using the sound of a whistle to mark a desired behavior, followed by a food reward. For such demonstrations a new behavior was chosen for training in each of four daily public demonstrations. Much to the surprise of both the trainers and the audience Malia offered an entirely new and dramatic behavior on her own: she swam rapidly and then rolled over on her back, lifted her tail vertically in the air, and coasted across the tank without propulsion. Malia continued to come up with new behaviors over the next few weeks, demonstrating some astonishing departures from normal dolphin behavior, including leaping upside down, swimming while revolving under water in corkscrew fashion, and, while upside down underwater, drawing long arcing lines in the silt on the tank bottom with the tip of her dorsal fin.

The philosopher Gregory Bateson, a resident scientist at the Oceanic Institute adjoining Sea Life Park, took an interest in this phenomenon and drew it to the attention of the Office of Naval Research, where studies of dolphin cognition were under way for possible military applications. A research grant made possible a repetition of the procedures that had resulted in spontaneous production of novel behaviors, using an experimentally naïve juvenile female rough-toothed dolphin named Hou. Extensive prior experience with the behavior of the dolphins enabled the trainers to identify and reinforce novel movements, and to decide during their occurrence whether or not the specific movement met the definition of novelty: "...any movement that was not part of the normal swimming action of the animal." (Pryor, Haag, & O'Reilly 1969, p. 654). Such movements were accompanied by a synchronous conditioned reinforcer (a short blast of a whistle), followed by delivery of the primary reinforcer, a small fish.

¹ The paper was titled *The Creative Porpoise* in deference to biologist Kenneth S. Norris, Ph.D., Sea Life Park advisor, who preferred this term for all small cetaceans. Today we would refer to this species as a dolphin, as the *true* porpoises are a separate group taxonomically.

Three observers watched each session: the trainer, an observer watching above water, and an observer viewing from below through the glass walls of the tank. The team was connected by headsets and microphones, allowing their comments and the trainer's whistle to be recorded on a single tape for the sessions. Sessions ranged from a few minutes to half an hour in length. Events were transcribed in 15 second intervals to create cumulative records for each of the 32 sessions.

Given that the research plan was based upon the unexpected novel behavior of a single well-trained dolphin, it is not surprising that with a new and less experienced subject the experimental plan of reinforcing a new type of response in each session was not met. For example, in order to maintain responding, it was sometimes necessary to include previously reinforced behavior within the session. In spite of the lack of experience on the part of the experimenters, the naïveté of the second dolphin, and the animal's tendency to lapse into repetitive circling when unreinforced, innovative behavior was established and brought under stimulus control.

A particularly striking finding of the Pryor et al. (1969) study was the sudden emergence of a burst of activity on the 16th day of training accompanied by a marked increased occurrence of novel behavior.

Hou became extremely active when the trainer appeared and immediately offered twisting breaches, landing on its belly and its back. It also began somersaulting on its long axis in mid-air . . . the flip occurred 44 times, intermingled with some of the previously reinforced responses, and with three other responses that had not been seen before: an upside-down tailslap, a sideswipe with the tail, and an aerial spin on the short axis of the body. (p. 657)

The burst in activity was accompanied by a dramatic increase in production of novel behaviors; not unlike the increase in choice accuracy following chance performance on a difficult choice discrimination task. According to Chase, Schupak, and Ploog (2012) it is during this *presolution period* that information gradually accumulates about the relationship between the sensations experienced, and the outcome of one's behavior. Hou's training ended after 33 sessions, when the behaviors had become so complex that it was difficult for the trainers to discriminate and describe them. While both animals were capable of meeting criteria there were individual differences. Hou offered responses within the normal observed behavioral range of the species. Malia offered responses differing from any behavior observed previously by these trainers in any dolphin species.

After the experimental sessions ended, Hou and Malia each continued to display the ability to generate novel behavior both within creativity training sessions and on their own.

In the following years Hou has learned to leap tank partitions to gain access to other porpoises, a skill very seldom developed by a captive porpoise. When a trainer was occupied at an adjoining porpoise tank, Malia jumped from the water, skidded across 6 ft of wet pavement, and tapped the trainer on the ankle with its rostrum or snout, a truly bizarre act for an entirely aquatic animal. (p. 661)

In the years following the initial publication, the first author and many colleagues in the animal training community have refined, replicated, and further developed techniques that facilitate the development of innovative behavior. Creativity training of both domestic and wild animals in confinement can reduce the stress levels associated with a monotonous or under-stimulating environment (Maple, 2007). These techniques are now in wide use by dog trainers, and by zoo and aquarium personnel across a wide variety of species (Pryor, 2009, pp. 82-86).

Meanwhile, very little has appeared in the research literature regarding the mechanisms underlying these procedures. An exception is an article by Schusterman and Reichmuth (2008) describing the training of two walruses to produce novel vocalizations. They noted that as the walruses became more experienced in the task they “often shifted their sound emissions quite rapidly, and appeared to vary their behavior in order to explore which sound(s) would be selected for reinforcement” (p. 324).

The Event Marker

A crucial factor in training the dolphins in the Pryor et al. (1969) study was the use of an event marker, in this case a brief whistle, presented during (not after) the selected behavior, to isolate the specific action selected for reinforcement. The event marker acquires the properties of a conditioned reinforcer through association with the unconditioned reinforcer. Prior to its use in training its reinforcing properties are established through a short series of such pairings. However, unlike the researchers’ use of conditioned reinforcers to train new behaviors in the absence of the unconditioned reinforcer, in marker-based training the unconditioned reinforcer is always presented after the behavior to be reinforced has been marked. This procedure results in a strong and a consistent association between the event marker and positive reinforcement, an association free from ambiguous and sometimes negative associations as might be true for the use of the voice. Once established as a conditioned reinforcer, the same event marker can be used in training other behaviors, thus developing a wide variety of behaviors in a single learner. Each such behavior may then be brought under stimulus control for performance or other purposes.

Any distinctive stimulus can be an event marker. The stimulus chosen for this function should be brief, initially neutral, distinct from other stimuli in the environment, easily delivered, and perceivable by both trainer and learner. That the marker should be precise, initially neutral, and different from other stimuli in the environment was experimentally demonstrated by Wood (2006) who compared the sound of the clicker to the spoken word *good*, a less precise and far from unique stimulus, and one that may also have previous negative associations.

As early as 1963, most marine mammal trainers had settled on the whistle as the marker of choice, as it met the criteria for an effective marker and was audible both in the air and under the water. The clicker has since become the event marker of choice for dog and horse trainers as well as in training zoo animals. A light (such as the blink of a flashlight) is now commonly used for capturing and shaping behavior in fish. Touch, vibration, and even scent are also possibilities if called for by the sensory capacities of the learner. The clicker is being used in some veterinary hospitals and also some laboratories. An example of laboratory use of a clicker is in the study of free movement in cats at the Barrow Neurological Institute in Phoenix, Arizona.

Marker-based training was first explicated in a book for the general public by the senior author (Pryor, 1984, 1999). The specific use of a clicker as an event marker was publicly introduced by Pryor and three other trainers at the 1992 annual meeting of the Association for Behavior Analysis International. Fueled by the Internet, seminars, and an expanding popular literature, the technology spread rapidly among dog trainers and zoo keepers (Pryor, 2009, Pryor & Ramirez, 2014). These procedures have also come into use in a range of human applications, including sports and special needs education (Persicke & Adams, 2013), and show promise of being helpful in the development of basic skills in the severely autistic (Gabler, 2013). When used to shape behavior in adults and children, the click is sometimes labeled a TAG (Teaching with Acoustical Guidance) to reduce the association with clicks used in dog training.

Trainers generally recognize the event marker as being particularly effective in the following circumstances: capturing new or uncommon behavior; reinforcing small variations in a movement in order to

raise criteria incrementally in shaping a new behavior; and isolating and strengthening small elements of larger movements, such as a gymnast pointing the toes in a handstand, or a show horse pointing the ears forward while moving under saddle.

The marker is valuable in reinforcing a behavior at a distance, when the event occurs away from the source of the food or other primary reinforcers: for example, while a trained dolphin is searching for objects on the sea floor (Pryor, 1973), or a dog is running over obstacles in an agility course; in this circumstance the marker is sometimes referred to as a bridging stimulus.

Extensive work by the animal training community has demonstrated the efficacy of event marking and identified conditions under which it is most effective; however, the focus of the trainer is on the effectiveness of the procedures, rather than investigation into the underlying mechanisms. As a result, very little of this work has been published. Fortunately, these issues have not escaped the attention of those interested in basic research. The following is a brief summary of some of the current understanding of event marking that has been experimentally validated through laboratory research.

Experimental Verification of Efficacy of Event Marking.

An important characteristic of the marker is that it serves to bridge the temporal gap between the behavior to be conditioned and the delivery of reinforcement. Grice (1948) found that even a short delay of 0.5 seconds between the response and reinforcer may result in little or no learning. The improvement in task acquisition by marking the behavior to be conditioned prior to delivery of the primary reinforcer has been clearly shown in a series of articles by Lieberman and associates (Lieberman, McIntosh, & Thomas, 1979; Thomas, Lieberman, McIntosh, & Ronaldson, 1983) using rats and pigeons in two-choice situations. These marking experiments also validated the clicker trainers' description of the qualities of an effective marker. Laboratory work led Thomas et al. (1983) to conclude that a marker is effective in enhancing learning only when it is temporally contiguous to the to-be-remembered event. They found that "the critical feature enabling rats to form associations over a delay in these maze experiments was the occurrence of some brief, salient event (handling, white noise, a light) immediately after the rat had chosen one of the two alleys." (p. 126).

Work on sign tracking (e.g., Hearst & Jenkins, 1974) provides evidence that a neutral stimulus, e.g. a clicker, may be responded to in a manner similar to that associated with the unconditioned stimulus. These findings are supported by accumulating evidence from neurophysiology, for example, that a neutral stimulus may acquire incentive properties similar to that of the primary reinforcer (Flagel et al., 2011; Wise, 2004). Therefore using the marker to precisely identify the behavior required may be thought of as equivalent to using an immediate presentation of the reinforcer, something often difficult or impossible to do in training complex behaviors.

The use of an event marker is particularly relevant in training behaviors that are very different from those associated with the reinforcer. That behavior varies systematically and sequentially in anticipation of food delivery was shown clearly by Staddon and Simmelhag (1971). Anticipation of eating, with its associated movements, if incompatible with the movements to be trained, blurs the distinction. Thus one of the possible reasons for the efficacy of the marker is that it has the property of separating the behavior associated with the experience of sensing and remembering from the behavior associated with the experience of consumption.

Innovation and Behavioral Variability

Innovation or creativity is not exclusive to humans, nor does it require the systematic training for novelty the dolphins in the Pryor et al. (1969) study received. Biologists have identified innovative behavior in many species in the wild (Reader & Laland, 2003), and across many taxa, including fish (Laland & Van Bergen, 2003). Innovative behavior may turn out to be a widespread adaptive trait.

That it does not require a trainer who reinforces novelty is beautifully documented in the study by Russon, Kunoro, Firisa, and Handayani (2010) of innovative ways of using water by rehabilitated orangutans living in a natural habitat. Twenty months and 1,397 hours of focal observations were used to document water activity of 41 orangutans. Innovations, a subset of 18 out of 44 variants which showed a functional and behavioral gain of water use, were generally observed as small changes emerging from prior behaviors through social facilitation, accidental occurrence, application of old behaviors to new ends, or worked out independently by individuals. In many cases these processes occurred in combination.

Although Russon et al. (2010) did not mention observing increases in behavioral variability, it is quite likely that, given the complexity of their environment, changes occurred under which old solutions were no longer effective. It is well established that behavior tends to become more variable during reduced reinforcement (e.g., Antonitis, 1951, Eckerman & Lanson, 1969; Gharib, Gade, & Roberts, 2004). Behavioral variability has the potential of generating new behaviors, some of which may be novel. Thus, this clearly plays a part in creativity.

Behavioral variability during non-reinforcement of previously reinforced behaviors was clearly seen in the behavior of the dolphin Hou (Pryor et al., 1969) early in the experiment, when the previously reinforced behavior of porpoising (leaping smoothly out then in the water) was not reinforced.

...the animal rapidly developed a behavior pattern of porpoising in front of the trainer, entering the water in an inverted position, turning right side up, swimming in a large circle, and returning to porpoise in front of the trainer again. It did this 25 times without interruption over a period of 12.5 min. Finally, it stopped and laid its head against the pool edge at the trainer's feet. (p. 655)

This new behavior was then reinforced. Although, as this example illustrates, non-reinforcement of previously learned behavior was effective in generating a new behavior, alternative means exist for generating novel behaviors without discontinuing reinforcement. For example, in order to stimulate exploratory behavior one may introduce a novel object and then mark and reinforce, in rapid succession, any response to the new object, such as looking at it, approaching, smelling, touching, climbing on, into, or under the object, manipulating it, and so on. Hiding food items around the training area and marking and reinforcing behavior related both to searching and to finding the hidden food is another tool for increasing opportunities for reinforcing varied behavior. Standard procedures used to optimize learning under both applied and laboratory conditions, such as adjustment of session length to avoid fatigue or satiation, or conducting training in a specific location under defined conditions, are also conducive to generation of novel behavior. The latter is particularly important in demanding tasks, such as one in which innovation is required.

Laboratory Studies of Reinforced Variability

The work of Allen Neuringer (2004) and his colleagues (e.g., Neuringer, Deiss, & Olson, 2000; Neuringer & Jensen, 2010), on behavioral variability has provided extensive data on the generation and stimulus control of sequences of behavior, varying from repetition to randomness, by pigeons, rats and

humans. Although much has been learned by the examination of the sequences of bar and key pressing, in such experiments the degree of behavioral variability is limited to whatever is behaviorally possible when the required behavior is simply closing a switch. This restriction was somewhat relaxed by Stahlman and Blaisdell's (2011a) use of a touch screen to record variations in the horizontal and vertical coordinates of pigeons' pecks. Stahlman and Blaisdell (2011b) extended this further to an examination of variability of search location and search frequency, the paths taken by rats seeking hidden food. However, even touch screens or a platform to be traversed cannot capture complex motor actions such as the new patterns of movements exhibited by the Pryor et al. (1969) dolphins, by many other species reinforced for creativity by trainers, and, most obviously, by the conditions present in the natural environment, as shown by Russon et al. (2010).

Important progress towards the understanding of more complex behavior has been made by Manabe, Staddon, and Cleaveland (1997). By focusing on vocalization of budgerigars, a behavior that could be computer processed and thus identified and reinforced automatically, they were able to show that budgerigars can generate novel call patterns when required to do so to obtain food. The birds modified and recombined existing calls when reinforcement became contingent upon production of a call sufficiently different from up to the last three calls. One might say that this was not truly novel behavior, but simply recombination of previously learned behavior, had it not been for one bird that met the requirement—varying the calls—by pecking on the wall of the chamber. The Manabe et al. (1997) experiment, as well as the naturalistic experiments of Russon et al. (2010), show that novel behavior can emerge, even without specific reinforcement of novelty, under conditions in which only behavioral variability is required.

Effects of Non-Reinforcement

Much of the success of marker based training rests on the consistent use of positive reinforcement and the avoidance of aversive procedures such as punishment. That the removal of reinforcement can provoke new and varied behavioral responses is well known (Antonitis, 1951, Gharib et al., 2004). However non-reinforcement carried out to extinction is clearly aversive. When it occurs in training, the animal may abandon its efforts entirely. Negative states of affect are best avoided in a training situation that depends on the animal's willing participation, in which the animal may be expected to learn and even initiate many behaviors over a long period of time.

As the work on behavioral variability shows, non-reinforcement plays an essential role in discriminating between what does and does not work. As such, both the presence and the absence of the marker and reinforcer provide information and facilitate new patterns of behavior. Systematic research is necessary to determine the optimal balance between the positive effects of the marker and the stress associated with non-reinforcement in the emergence of new behaviors. Particularly relevant here is that extinction-induced emotional states may inhibit what we identify as creativity. By maintaining a high reinforcement rate over a wide range of behaviors one also maintains the subject's positive state of affect, facilitating the development of exploratory and innovative behavior.

In a complex study of rewards and creative performance in humans, Byron and Khazanchi (2012) concluded that creativity-contingent rewards have a stronger positive effect on creativity than performance-contingent rewards, especially for complex tasks in which frustration may occur. Although the human research is vastly more complex than the demonstrations of creativity in animals discussed here, creativity training in animals appears to reflect the effects of conditions that may increase or decrease creativity-related behavior in humans as well.

Conclusion

The long and distinguished history of research into reinforcement that has led to the development of the current marker based technology is enriching our awareness of animal internal states of affect, innate behavior, and cognition. However, as the present review of the literature suggests, important issues relevant to the generation of innovative behavior remain unexplored. The principles underlying acquisition of variability and novel behavior could contribute to the work of behavioral ecologists and cognitive psychologists, as well as the work of neuroscientists studying underlying physiological processes. This cross-fertilization of disciplines could lead to productive outcomes both in animal and in human research, with the prospect of new and valuable applications emerging in the future. A deeper understanding of these issues in animals may enhance the understanding of conditions conducive to innovation in humans as well. The current paucity of published research into development of creativity in animals constitutes a rich domain for future exploration and research.

References

- Antonitis, J. J. (1951). Response variability in the white rat during, conditioning, extinction and reconditioning. *Journal of Experimental Psychology*, 42, 273-281.
- Byron, K., & Khazanchi, S. (2012). Rewards and creative performance: A meta-analytic test of theoretically derived hypotheses. *Psychology Bulletin*, 138, 809-830.
- Chase, S., Schupak, C., & Ploog, B. O. (2012). Attention, the presolution period, and choice accuracy in pigeons. *Behavioural Processes*, 89, 225-231.
- Eckerman, D. A., & Lanson, R. N. (1969). Variability of response location for pigeons responding under continuous reinforcement, intermittent reinforcement, and extinction. *Journal of the Experimental Analysis of Behavior*, 12, 73-80.
- Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature*, 469, 53-59.
- Gabler, M. (2013). *Chaos to calm: Discovering solutions for living with autism*. Campbellville, Ontario, Canada. TAGteach Press.
- Gharib, A., Gade, C., & Roberts, S. (2004). Control of variation by reward probability. *Journal of Experimental Psychology: Animal Behavior Processes*, 30, 271-282.
- Grice, G. R. (1948). The relation of secondary reinforcement to delayed reward in visual discrimination learning. *Journal of Experimental Psychology*, 38, 1-16.
- Hearst, E., & Jenkins, H. (1974). Sign-tracking: The stimulus reinforcer relation and directed action. *Monograph of the Psychonomic Society*.
- Laland, K. N., & Van Bergen, Y. (2003). Experimental studies of innovation in the guppy. In S. M. Reader & K. N. Laland (Eds.), *Animal innovation*. Oxford, UK: Oxford University Press.
- Lieberman, D. A., McIntosh, D. C., & Thomas, G. V. (1979). Learning when reward is delayed: A marking hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes*, 5, 224-242.
- Manabe, K., Staddon, J. E. R., & Cleaveland, J. M. (1997). Control of vocal repertoire by reward in budgerigars. *Journal of Comparative Psychology*, 111, 50-62.
- Maple, T. L. (2007). Towards a science of welfare for animals in the zoo. *Journal of Applied Welfare Science*, 10, 63-70.
- Neuringer, A. (2004). Reinforced variability in animals and people: Implications for adaptive action. *American Psychologist*, 59, 891-906.
- Neuringer, A., Deiss, C., & Olson, G. (2000). Reinforced variability and operant learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 26, 98-111.
- Neuringer, A., & Jensen, G. (2010). Operant variability and voluntary action. *Psychological Review*, 117, 972-

993.

- Persicke, A., Jackson, M. & Adams, A.N. (2013). Brief report: An evaluation of TAGteach components to decrease toe-walking in a four year old child with autism. *Journal of Autism Developmental Disorders*, 44, 965-968.
- Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior*, 12, 653-661.
- Pryor, K. (1973). *Lads before the wind. Adventures in dolphin training*. New York, NY: Harper & Row.
- Pryor, K. (1999). *Don't shoot the dog! The new art of teaching and training*. Rev. Ed. New York, NY: Bantam Books.
- Pryor, K. (2009). *Reaching the animal mind* (pp. 82-87, 90-95). New York, NY: Scribner.
- Pryor, K., & Ramirez, K. R. (2014). Modern animal training: A transformative technology. In F. McSweeney & E. Murphy (Eds.), *A handbook of operant and classical conditioning*. New York, NY: Wiley and Blackwell.
- Reader, S. M., & Laland, K. N. (Eds.). (2003). *Animal innovation*. Oxford, UK: Oxford University Press.
- Russon, A. E., Purwo, K., Ferisa, A., & Handayani, D. P. (2010). How orangutans (*Pongo pygmaeus*) innovate for water. *Journal of Comparative Psychology*, 124, 14-28.
- Schusterman, R. J., & Reichmuth, C. (2008). Novel sound production through contingency learning in the Pacific walrus (*Odebenus rosmarus divergens*). *Animal Cognition*, 11, 319-327.
- Staddon, J. E. R., & Simmelhag, V. L. (1971). The superstition experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review*, 78, 3-43.
- Stahlman, W. D., & Blaisdell, A. P. (2011a). Reward probability and variability of foraging behavior in rats. *International Journal for Comparative Psychology*, 24, 168-176.
- Stahlman, W. D. & Blaisdell, A. P. (2011b). The modulation of operant variation by the probability, magnitude, and delay of reinforcement. *Learning and Motivation*, 42, 221-236.
- Thomas, G. V., Lieberman, D. A., McIntosh, D. C., & Ronaldson, P. (1983). The role of marking when reward is delayed. *Journal of Experimental Psychology: Animal Behavior Processes*, 9, 401-411.
- Wise, R. A. (2004). Dopamine, learning and motivation, *Nature Reviews Neuroscience*, 5, 483-494.
- Wood, L. A. (2006). *An analysis of the efficacy of bridging stimuli: Comparing the clicker to a verbal bridge* (Unpublished Master's thesis), Hunter College, CUNY, NY.

Financial Support: N/A

Conflict of Interest: The author declared no conflict of interest.

Submitted: October 28th, 2013

Resubmitted: March 24th, 2014

Accepted: April 9th, 2014